

```

1 Discussion 2020-07-29: String suggestion problem.
2 -----
3
4 Minu
5 -----
6     >> Inverted Index
7     >> Evaluation metric ??
8     >> Correct, Correct word not predicted ??
9     >> ?? if else expression ???
10
11 Binod
12 =====
13     >> Count vectorizer: Problem
14     >> how to do it?
15
16     (ngram generation completed)
17     >> bigram
18     >> tf-idf concept ??
19
20 Shankar
21 =====
22     >> zero remove ??
23     >> clojure variable ??
24     >>
25
26 Astha
27 =====
28     >> Algorithm: inverted index, vectorizer
29     >> Suggestion adequacy
30
31 Data generation
32 -----
33
34 -----
35 Removal of if-else expression
36 -----
37 Collection: you need to convert them.
38 [0, 1, 2, 6, 7] => apply isOdd
39
40 odd numbers
41
42 [] => ?
43
44 7 => [7] => isEmpty?, isEmpty?
45
46 filter, (fn[] ()), vector/list => final output.
47
48 -----
49 Zero remove(default value)
50 Collection => getOrElse(key, defaultValue)
51 -----
52
53 clojure variable
54 -----
55 const x => 100
56 reduce(Map()) {
57     .....
58     (* x variable)
59 }
60 -----
61
62 Data generation
63 -----
64 for()

```

```

65 first 100 prime numbers
66 -----
67 >> infinite seq (lazy).fold((result [])).filter..... take(100)
68
69 -----
70 tail recursion
71 -----
72 f(10), 1
73 f(9), 10
74 f(8), 90
75 f(7), 720
76
77 ...
78 f(1), result.
79 take result.
80
81 -----
82 for (i: 1 to 10) {
83     i * result.
84 }
85
86 expansion => first expand, and then evaluate.
87
88 generate, accumulate.
89
90 0, negative number.
91 -----
92 cond = [0, -1]
93
94 -9078803474589723498752093844759082347450982730948579082834750982734023445982734059823
95 239048523452345
96
97 count problem--
98 -----
99 [a, b, c, a, b, c, a]
100 -----
101 group by
102 -----
103 [a (a, a, a), b (b, b), c (c, c)]
104 map (count values)=> [a 3, b 2, c 2]
105 -----
106
107 Evaluation Metric
108 -----
109 Job 1: Evaluation Data
110 -----
111 word list
112 10000
113
114 test words
115 100
116 w1      => w3, w790, w678, test data
117 w100    =>
118
119 -----
120 w1      => system suggestion, w790, w3, w800
121
122 We can count hit and miss
123 -----
124 hit: 2, miss: 1, insertion 1, deletion ?
125
126 hit => true position...
127 miss => false negative
128 insertion => false positive

```

```
129  deletion => ...
130
131  accuracy: correct result, incorrect result
132  correct => 1500
133  incorrect => 500
134  total => 2000
135
136
137
138
139
140
```